Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(9): 17461–17474.
DOI:10.3934/math.2022962
Received: 01 May 2022
Revised: 17 July 2022
Accepted: 22 July 2022
Published: 28 July 2022

Research article

Multipoint flux mixed finite element method for parabolic optimal control problems

Tiantian Zhang[1], **Wenwen Xu**[1,*], **Xindong Li**[1] **and Yan Wang**[2]

[1] School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, Shandong, China

[2] College of Information Science and Engineering, Shandong Agricultural University, Taian 271018, Shandong, China

* **Correspondence:** Email: xuwenwen0922@163.com.

Abstract: In this paper, we research semi-discrete multipoint flux mixed finite element (MFMFE) method for parabolic optimal control problem (OCP). The state and co-state variables are approximated by the lowest order Brezzi-Douglas-Marini (BDM) mixed finite element (MFE) spaces and the control is approximated by piecewise constant. The advantage of this type of mixed element method is that it can decouple the state and adjoint state variables as cell-centered difference schemes rather than to solve saddle point algebraic equations. Error estimates and convergence orders are derived rigorously for state and control variables. Finally, a numerical example is given to confirm our theoretical analysis.

Keywords: multipoint flux mixed finite element; parabolic equation; optimal control problem; error estimate
Mathematics Subject Classification: 65N12, 65N15

1. Introduction

OCPs have a profound research background. They have been widely used in industrial process, transportation, economic management and other fields, and have been a great development in modern life. Undoubtedly, the finite element method (FEM) is commonly applied in computing OCP as well as there has been a lot of works [1–9]. The FEM for PDEs and OCPs are introduced systematically in references [10–12].

About solving the OCP of elliptic equation, there are plenty of works. In particular, here is an early article by Falk [2] on the prior error estimation of elliptic equations by applying the standard finite element method. Malanowshi [3] has established a priori error estimation for convex constrained optimal control using finite element approximations. Hou [5] also works out an article on error estimation and superconvergence of elliptic optimal control by applying MFE method. Guo [13] innovatively used a split positive definite MFE method to deal with the convex OCP of elliptic equation. Liu and Chen [14] use MFE method to calculate OCP and obtain a posteriori error estimation. However, there are not many studies on parabolic equations. For convex polygon region parabolic OCP, there is a prior error for the FEM approximation, see Shakya [15]. Xing, Lu and Chen [16, 17] have obtained L^2 error and a priori error in allusion to parabolic OCPs by using MFE method.

In this paper, we will adopt semi-discrete MFMFE method to develop L^2 error of parabolic control problem. MFMFE method is based on the lowest order BDM_1 MFE space. By adopting special quadrature rules, it can not only allow local flux elimination, but also can be formed the cell-centered system for state varibles and co-state varibles, also this system is symmetrically positive definite, see [18–21]. That avoids the problem of solving saddle-point algebraic systems required by MFE method.

Other contents are arranged. The MFMFE method is introduced in Section 2. The L^2 error is described in Section 3. A numerical example is used to verify the correctness of theoretical analysis in Section 4. We summarize the paper in Section 5. We specifically study the OCP for the state variable ϱ, q and the control ς:

$$\min_{\varsigma\in K}\left\{\frac{1}{2}\int_0^T (\|\varrho-\varrho_d\|^2+\|q-q_d\|^2+\|\varsigma\|^2)dt\right\}, \tag{1.1}$$

according with the following conditions

$$\begin{cases} q_t+div\varrho=\varsigma, \varrho=-B\nabla q, x\in\Omega,\\ q(x,t)=0, x\in\partial\Omega, t\in J,\\ q(x,0)=q_0(x), x\in\Omega, \end{cases} \tag{1.2}$$

where $J=[0,T]$. Here, we define $\Omega\subset R^2$ and it is a bounded convex domain with C^2 boundary. As well as $B(x)$ is a symmetric positive definite matrix. K represents the admissible set of the control variable, given by

$$K=\left\{\tilde{\varsigma}\in L^2(\Omega):\alpha\le\tilde{\varsigma}\le\beta\right\}, \tag{1.3}$$

where α and β are given functions. Next, we give the notation of the space that we need to use. First, $W^{m,p}(\Omega)$ is defined by the Sobolev spaces on Ω. And it has two related norms, which are defined by

$$\|\phi\|_{m,p}=\sum_{|\alpha|\le m}\|D^\alpha\phi\|^p_{L^p(\Omega)},$$

$$|\phi|_{m,p}=\sum_{|\alpha|=m}\|D^\alpha\phi\|^p_{L^p(\Omega)}.$$

Here, $W_0^{m,p}(\Omega)$ means equal to 0 at the boundary. We define that $H^m(\Omega)=W^{m,2}(\Omega)$, $H_0^m(\Omega)=W_0^{m,2}(\Omega)$, and $\|\cdot\|_m=\|\cdot\|_{m,2}$, $\|\cdot\|=\|\cdot\|_{0,2}$. We denote by $L^s(J;W^{m,p}(\Omega))$ the Banach space of all L^s integrable functions from J into $W^{m,p}(\Omega)$ with norm

$$\|\theta\|_{L^s(J;W^{m,p}(\Omega))}=(\int_0^T\|\theta\|^s_{W^{m,p}(\Omega)}dt)^{\frac{1}{s}},$$

here $s\in[1,\infty)$.

2. The multipoint flux method

This section mainly studies MFMFE method approximation problem (1.1) and (1.2). There are

$$V = H(div) = \left\{\varpi \in (L^2(\Omega))^2, div\varpi \in L^2(\Omega)\right\},$$

$$W \in L^2(\Omega),$$

the weak formulation of (1.1) and (1.2) can be written as: find $(\varrho, q, \varsigma) \in V \times W \times K$ such that

$$\min_{\varsigma\in K}\left\{\frac{1}{2}\int_0^T (\|\varrho-\varrho_d\|^2 + \|q-q_d\|^2 + \|\varsigma\|^2)dt\right\}, \tag{2.1}$$

$$(B^{-1}\varrho, \mathbf{v}) - (q, div\mathbf{v}) = 0, \forall \mathbf{v} \in V, \tag{2.2}$$

$$(q_t, \pi) + (div\varrho, \pi) = (\varsigma, \pi), \forall \pi \in W, \tag{2.3}$$

where $(\cdot)$ represents the inner product in $L^2(\Omega)$ or $(L^2(\Omega))^2$. The convex control problems (2.1)–(2.3) has a unique solution (ϱ, q, ς). In addition the triplet (ϱ, q, ς) is the solution to (2.1)–(2.3) if and only if there is a co-state $(g, r) \in V \times W$ such that $(\varrho, q, g, r, \varsigma)$ satisfies the optimality conditions for $t \in J$ [14–16]:

$$(B^{-1}\varrho, \mathbf{v}) - (q, div\mathbf{v}) = 0, \forall \mathbf{v} \in V, \tag{2.4}$$

$$(q_t, \pi) + (div\varrho, \pi) = (\varsigma, \pi), \forall \pi \in W, \tag{2.5}$$

$$(B^{-1}g, \mathbf{v}) - (r, div\mathbf{v}) = -(\varrho - \varrho_d, \mathbf{v}), \forall \mathbf{v} \in V, \tag{2.6}$$

$$-(r_t, \pi) + (divg, \pi) = (q - q_d, \pi), \forall \pi \in W, \tag{2.7}$$

$$(r + \varsigma, \tilde{\varsigma} - \varsigma) \geq 0, \forall \tilde{\varsigma} \in K. \tag{2.8}$$

We are going to give you our own method. MFMFE method is adopted, which considers partition, mapping, BDM_1 space and special quadrature rules as in the article [20]. Then we display some theories that are important to us. The spaces on Γ_h are made up of

$$V_h = \left\{v \in V : v|_E \leftrightarrow \hat{v}, \hat{v} \in \hat{V}(\hat{E}), \forall E \in \Gamma^h\right\}, \tag{2.9}$$

$$W_h = \left\{\pi \in W : \pi|_E \leftrightarrow \hat{\pi}, \hat{\pi} \in \hat{W}(\hat{E}), \forall E \in \Gamma^h\right\}, \tag{2.10}$$

$$K_h = \left\{\tilde{\varsigma} \in W_h : \tilde{\varsigma}|_E \in ((\alpha)_E, (\beta)_E), \forall E \in \Gamma^h\right\}, \tag{2.11}$$

where $v|_E \leftrightarrow \hat{v}$ is a mapping from the subdivision cell E to the reference cell $\hat{E} = [0, 1] \times [0, 1]$, and V_h, W_h are BDM_1 spaces. In addition,

$$(\alpha)_E = \frac{1}{|\mathrm{E}|}\int_E \alpha(x, t)dx,$$

$$(\beta)_E = \frac{1}{|\mathrm{E}|}\int_E \beta(x, t)dx,$$

for $r, v \in V_h$, we give the definition of global quadrature rule [21] which is limited to $\hat{E}$,

$$(B^{-1}r, v)_Q = \sum_{E\in\Gamma^h}(B^{-1}r, v)_{Q,E}.$$

According to (2.9) and (2.10), we can get

$$\int_E B^{-1}r\cdot vdx = \int_{\hat{E}}\hat{B}^{-1}\frac{1}{J_E}DF_E\hat{r}\cdot\frac{1}{J_E}DF_E\hat{v}J_Ed\hat{x} = \int_{\hat{E}}\frac{1}{J_E}DF_E^T\hat{B}^{-1}DF_E\hat{r}\hat{v}d\hat{x} = \int_E\mathscr{B}^{-1}\hat{r}\hat{v}d\hat{x},$$

where $\mathscr{B} = J_EDF_E^{-1}\hat{B}(DF_E^{-1})^T$ and $F_E : \hat{E} \leftrightarrow E$ is a bijection mapping. As well as DF_E is a Jacobi matrix and J_E is a Jacobian. Other details see references [21]. Represents the quadrature error of element through

$$\sigma_E(B^{-1}r, v) = (B^{-1}r, v)_E - (B^{-1}r, v)_{Q,E},$$

as well as the global quadrature error is given through $\sigma(B^{-1}r, v)|_E = \sigma_E(B^{-1}r, v)$.
Now, we introduce two known conclusions from the reference [21].

Lemma 2.1. *For one constant M, it is nothing to do with h, such that*

$$(B^{-1}r, r)_Q \geq M\|r\|^2, \forall r \in V_h.$$

Lemma 2.2. *For one constant M, it is nothing to do with h, such that*

$$\sigma(B^{-1}\Pi_hv, \pi) \leq Mh\|v\|_1\|\pi\|, \forall\pi \in W_h.$$

See the literature [20, 21] for more details on quadrature rules. Then the MFMFE method approximation for the problem (2.1)–(2.3) is to find $(\varrho_h, q_h, \varsigma_h) \in V_h \times W_h \times K_h$ such that

$$\min_{\varsigma_h\in K_h}\left\{\frac{1}{2}\int_0^T(\|\varrho_h-\varrho_d\|^2 + \|q_h-q_d\|^2 + \|\varsigma_h\|^2)dt\right\}, \tag{2.12}$$

$$(B^{-1}\varrho_h, \mathbf{v}_h)_Q - (q_h, div\mathbf{v}_h) = 0, \forall\mathbf{v}_h \in V_h, \tag{2.13}$$

$$(q_{ht}, \pi_h) + (div\varrho_h, \pi_h) = (\varsigma_h, \pi_h), \forall\pi_h \in W_h. \tag{2.14}$$

The control problem (2.12)–(2.14) has a unique solution $(\varrho_h, q_h, \varsigma_h)$ and a triplet $(\varrho_h, q_h, \varsigma_h)$ is the solution of (2.12)–(2.14) if and only if there exists a costate $(g_h, r_h) \in V_h \times W_h$ so that $(\varrho_h, q_h, \varsigma_h, g_h, r_h)$ satisfies the optimality conditions:

$$(B^{-1}\varrho_h, \mathbf{v}_h)_Q - (q_h, div\mathbf{v}_h) = 0, \forall\mathbf{v}_h \in V_h, \tag{2.15}$$

$$(q_{ht}, \pi_h) + (div\varrho_h, \pi_h) = (\varsigma_h, \pi_h), \forall\pi_h \in W_h, \tag{2.16}$$

$$(B^{-1}g_h, \mathbf{v}_h)_Q - (r_h, div\mathbf{v}_h) = -(\varrho_h - \varrho_d, \mathbf{v}_h), \forall\mathbf{v}_h \in V_h, \tag{2.17}$$

$$-(r_{ht}, \pi_h) + (divg_h, \pi_h) = (q_h - q_d, \pi_h), \forall\pi_h \in W_h, \tag{2.18}$$

$$(r_h + \varsigma_h, \tilde{\varsigma}_h - \varsigma_h) \geq 0, \forall\tilde{\varsigma}_h \in K_h. \tag{2.19}$$

Next, we will take advantage of a few indirect variables. We give the definition of state solution $(\varrho(\tilde{\varsigma}), q(\tilde{\varsigma}), g(\tilde{\varsigma}), r(\tilde{\varsigma}))$, which $\forall\tilde{\varsigma} \in K$ are

$$(B^{-1}\varrho(\tilde{\varsigma}), \mathbf{v}) - (q(\tilde{\varsigma}), div\mathbf{v}) = 0, \forall\mathbf{v} \in V, \tag{2.20}$$

$$(q_t(\tilde{\varsigma}),\pi)+(div\varrho(\tilde{\varsigma}),\pi)=(\tilde{\varsigma},\pi),\forall\pi\in W, \tag{2.21}$$
$$(B^{-1}g(\tilde{\varsigma}),\mathbf{v})-(r(\tilde{\varsigma}),div\mathbf{v})=-(\varrho(\tilde{\varsigma})-\varrho_d,\mathbf{v}),\forall\mathbf{v}\in V, \tag{2.22}$$
$$-(r_t(\tilde{\varsigma}),\pi)+(divg(\tilde{\varsigma}),\pi)=(q(\tilde{\varsigma})-q_d,\pi),\forall\pi\in W. \tag{2.23}$$

We give the discrete state solution $(\varrho_h(\tilde{\varsigma}),q_h(\tilde{\varsigma}),g_h(\tilde{\varsigma}),r_h(\tilde{\varsigma}))$, which $\forall\tilde{\varsigma}\in K$ are

$$(B^{-1}\varrho_h(\tilde{\varsigma}),\mathbf{v}_h)-(q_h(\tilde{\varsigma}),div\mathbf{v}_h)=0,\forall\mathbf{v}_h\in V_h, \tag{2.24}$$
$$(q_{ht}(\tilde{\varsigma}),\pi_h)+(div\varrho_h(\tilde{\varsigma}),\pi_h)=(\tilde{\varsigma},\pi_h),\forall\pi_h\in W_h, \tag{2.25}$$
$$(B^{-1}g_h(\tilde{\varsigma}),\mathbf{v}_h)-(r_h(\tilde{\varsigma}),div\mathbf{v}_h)=-(\varrho_h(\tilde{\varsigma})-\varrho_d,\mathbf{v}_h),\forall\mathbf{v}_h\in V_h, \tag{2.26}$$
$$-(r_{ht}(\tilde{\varsigma}),\pi_h)+(divg_h(\tilde{\varsigma}),\pi_h)=(q_h(\tilde{\varsigma})-q_d,\pi_h),\forall\pi_h\in W_h. \tag{2.27}$$

3. Error analysis

Now, we estimate the error between the exact solution $(\varrho,q,g,r,\varsigma)$ and its approximation $(\varrho_h,q_h,g_h,r_h,\varsigma_h)$. First, we give the definition [16] of the standard L^2 projection $P_h:W\to W_h,\forall\theta\in W$, that is

$$(\theta-P_h\theta,\pi_h)=0,\forall\pi_h\in W_h. \tag{3.1}$$

We introduce the Fortin projection $\Pi_h:V\to V_h$, $\forall\mu\in V$, that is

$$(div(\mu-\Pi_h\mu),\pi_h)=0,\forall\pi_h\in W_h. \tag{3.2}$$

So we get certain projection relationships are as follows:

$$\|\theta-\Pi_h\theta\|\le Mh|\theta|_1,\theta\in(H^1(\Omega))^2, \tag{3.3}$$

$$\|div(\theta-\Pi_h\theta)\|\le Mh|div\theta|_1,div\theta\in H^1(\Omega), \tag{3.4}$$

$$\|\pi-P_h\pi\|_{-s}\le Mh^{1+s}|\pi|_1,\ s=0,1,\pi\in H^1(\Omega). \tag{3.5}$$

We review some existing conclusions in reference [16], which will be important for our next analysis.

Lemma 3.1. *For $M>0$, which is a constant and independent of h, we have*

$$\|\varrho-\varrho_h(\varsigma)\|_{L^2(J;L^2(\Omega))}+\|q-q_h(\varsigma)\|_{L^\infty(J;L^2(\Omega))}\le Mh, \tag{3.6}$$
$$\|g-g_h(\varsigma)\|_{L^2(J;L^2(\Omega))}+\|r-r_h(\varsigma)\|_{L^\infty(J;L^2(\Omega))}\le Mh. \tag{3.7}$$

Now, we just need to give the error estimates $\|\varrho_h(\varsigma)-\varrho_h\|_{L^2(J;L^2(\Omega))}$, $\|q_h(\varsigma)-q_h\|_{L^\infty(J;L^2(\Omega))}$, $\|g_h(\varsigma)-g_h\|_{L^2(J;L^2(\Omega))}$, $\|r_h(\varsigma)-r_h\|_{L^\infty(J;L^2(\Omega))}$.

Theorem 3.1. *For $M>0$, which is a constant and independent of h, we have*

$$\|\varrho_h(\varsigma)-\varrho_h\|_{L^2(J;L^2(\Omega))}+\|q_h(\varsigma)-q_h\|_{L^\infty(J;L^2(\Omega))}\le M\|\varsigma-\varsigma_h\|_{L^2(J;L^2(\Omega))}+Mh, \tag{3.8}$$
$$\|g_h(\varsigma)-g_h\|_{L^2(J;L^2(\Omega))}+\|r_h(\varsigma)-r_h\|_{L^\infty(J;L^2(\Omega))}\le M\|\varsigma-\varsigma_h\|_{L^2(J;L^2(\Omega))}+Mh. \tag{3.9}$$

Proof. First, let $\tilde{\varsigma} = \varsigma$ in (2.24)–(2.27), we will have the relations for intermediate solution $(\varrho_h(\varsigma), q_h(\varsigma), r_h(\varsigma), g_h(\varsigma))$.

$$(B^{-1}\varrho_h(\varsigma), \mathbf{v}_h) - (q_h(\varsigma), div\mathbf{v}_h) = 0, \forall \mathbf{v}_h \in V_h, \tag{3.10}$$

$$(q_{ht}(\varsigma), \pi_h) + (div\varrho_h(\varsigma), \pi_h) = (\varsigma, \pi_h), \forall \pi_h \in W_h, \tag{3.11}$$

$$(B^{-1}g_h(\varsigma), \mathbf{v}_h) - (r_h(\varsigma), div\mathbf{v}_h) = -(\varrho_h(\varsigma) - \varrho_d, \mathbf{v}_h), \forall \mathbf{v}_h \in V_h, \tag{3.12}$$

$$-(r_{ht}(\varsigma), \pi_h) + (divg_h(\varsigma), \pi_h) = (q_h(\varsigma) - q_d, \pi_h), \forall \pi_h \in W_h. \tag{3.13}$$

Then, we're going to take the difference between (2.15)–(2.18) and the Eqs (3.10)–(3.13) of the above intermediate solution. So we get error equations:

$$\begin{aligned}&(B^{-1}(\varrho_h(\varsigma) - \varrho), \mathbf{v}_h) + (B^{-1}(\varrho - \Pi_h\varrho), \mathbf{v}_h) + (B^{-1}(\Pi_h\varrho - \varrho_h), \mathbf{v}_h)_Q \\ &+\sigma(B^{-1}\Pi_h\varrho, \mathbf{v}_h) - (q_h(\varsigma) - q_h, div\mathbf{v}_h) = 0,\end{aligned} \tag{3.14}$$

$$(q_{ht}(\varsigma) - q_{ht}, \pi_h) + (div(\varrho_h(\varsigma) - \varrho), \pi_h) + (div(\Pi_h\varrho - \varrho_h), \pi_h) = (\varsigma - \varsigma_h, \pi_h), \tag{3.15}$$

$$\begin{aligned}&(B^{-1}(g_h(\varsigma) - g), \mathbf{v}_h) + (B^{-1}(g - \Pi_hg), \mathbf{v}_h) + (B^{-1}(\Pi_hg - g_h), \mathbf{v}_h)_Q \\ &+\sigma(B^{-1}\Pi_hg, \mathbf{v}_h) - (r_h(\varsigma) - r_h, div\mathbf{v}_h) = -(\varrho_h(\varsigma) - \varrho_h, \mathbf{v}_h),\end{aligned} \tag{3.16}$$

$$-(r_{ht}(\varsigma) - r_{ht}, \pi_h) + (div(g_h(\varsigma) - g), \pi_h) + (div(\Pi_hg - g_h), \pi_h) = (q_h(\varsigma) - q_h, \pi_h). \tag{3.17}$$

Next, the theorem is proved in two parts.

Part 1. Setting $v_h = \Pi_h\varrho - \varrho_h$ in (3.14), $\pi_h = q_h(\varsigma) - q_h$ in (3.15), combing these two equations, there is

$$\begin{aligned}&(B^{-1}(\Pi_h\varrho - \varrho_h), \Pi_h\varrho - \varrho_h)_Q + (q_{ht}(\varsigma) - q_{ht}, q_h(\varsigma) - q_h) \\ &= (\varsigma - \varsigma_h, q_h(\varsigma) - q_h) - (B^{-1}(\varrho_h(\varsigma) - \varrho), \Pi_h\varrho - \varrho_h) - (B^{-1}(\varrho - \Pi_h\varrho), \Pi_h\varrho - \varrho_h) \\ &-\sigma(B^{-1}\Pi_h\varrho, \Pi_h\varrho - \varrho_h) - (div(\varrho_h(\varsigma) - \varrho), q_h(\varsigma)).\end{aligned} \tag{3.18}$$

Note that

$$(B^{-1}(\Pi_h\varrho - \varrho_h), \Pi_h\varrho - \varrho_h)_Q \geq M||\Pi_h\varrho - \varrho_h||^2,$$

$$(q_{ht}(\varsigma) - q_{ht}, q_h(\varsigma) - q_h) = \frac{1}{2}\frac{d}{dt}||q_h(\varsigma) - q_h||^2.$$

Then, we can obtain

$$\begin{aligned}M||\Pi_h\varrho - \varrho_h||^2 + \frac{1}{2}\frac{d}{dt}||q_h(\varsigma) - q_h||^2 \leq\ & M||\varsigma - \varsigma_h||||q_h - q_h(\varsigma)|| + M||\varrho_h(\varsigma) - \varrho||||\Pi_h\varrho - \varrho_h|| \\ &+ M||\varrho - \Pi_h\varrho||||\Pi_h\varrho - \varrho_h|| + Mh||\varrho||_1||\Pi_h\varrho - \varrho_h|| \\ &+ M||div(\varrho_h(\varsigma) - \varrho)||||q_h(\varsigma) - q_h||,\end{aligned} \tag{3.19}$$

where we also used the inequality $\sigma(B^{-1}\Pi_hv, \pi) \leq Mh||v||_1||\pi||$. Next, we use the ε-Cauchy inequality, (3.3) and Lemma 3.1 to estimate (3.19),

$$M||\Pi_h\varrho - \varrho_h||^2 + \frac{1}{2}\frac{d}{dt}||q_h(\varsigma) - q_h||^2 \leq M||\varsigma - \varsigma_h||^2 + \varepsilon||\Pi_h\varrho - \varrho_h||^2 + \varepsilon||q_h(\varsigma) - q_h||^2 + Mh^2. \tag{3.20}$$

We integrate (3.20) over time as well as applying the Gronwall's lemma, where we will use $(q_h(\varsigma)-q_h)|_{t=0}=0$. So we get

$$\|\Pi_h\varrho-\varrho_h\|_{L^2(J;L^2(\Omega))}+\|q_h(\varsigma)-q_h\|_{L^\infty(J;L^2(\Omega))}\le M\|\varsigma-\varsigma_h\|_{L^2(J;L^2(\Omega))}+Mh. \tag{3.21}$$

Combining (3.21), (3.3) and (3.6), we will derive

$$\begin{aligned}&\|\varrho_h(\varsigma)-\varrho_h\|_{L^2(J;L^2(\Omega))}+\|q_h(\varsigma)-q_h\|_{L^\infty(J;L^2(\Omega))}\\&=\|\varrho_h(\varsigma)-\varrho\|_{L^2(J;L^2(\Omega))}+\|\varrho-\Pi_h\varrho\|_{L^2(J;L^2(\Omega))}\\&+\|\Pi_h\varrho-\varrho_h\|_{L^2(J;L^2(\Omega))}+\|q_h(\varsigma)-q_h\|_{L^\infty(J;L^2(\Omega))}\\&\le Mh+Mh+M\|\varsigma-\varsigma_h\|_{L^2(J;L^2(\Omega))}+Mh\\&\le M\|\varsigma-\varsigma_h\|_{L^2(J;L^2(\Omega))}+Mh.\end{aligned} \tag{3.22}$$

Part 2. Setting $v_h=\Pi_hg-g_h$ in (3.16), $\pi_h=r_h(\varsigma)-r_h$ in (3.17), combing these two equations, there is

$$\begin{aligned}&(B^{-1}(\Pi_hg-g_h),\Pi_hg-g_h)_Q-(r_{ht}(\varsigma)-r_{ht},r_h(\varsigma)-r_h)\\&=(q_h(\varsigma)-q_h,r_h(\varsigma)-r_h)-(\varrho_h(\varsigma)-\varrho_h,\Pi_hg-g_h)-(B^{-1}(g_h(\varsigma)-g),\Pi_hg-g_h)\\&-(B^{-1}(g-\Pi_hg),\Pi_hg-g_h)-\sigma(B^{-1}\Pi_hg,\Pi_hg-g_h)-(div(g_h(\varsigma)-g),r_h(\varsigma)-r_h).\end{aligned} \tag{3.23}$$

Note that

$$(B^{-1}(\Pi_hg-g_h),\Pi_hg-g_h)_Q\ge M\|\Pi_hg-g_h\|^2,$$

$$(r_{ht}(\varsigma)-r_{ht},r_h(\varsigma)-r_h)=\frac{1}{2}\frac{d}{dt}\|r_h(\varsigma)-r_h\|^2.$$

Then, using the same method as Part 1, we will get

$$\begin{aligned}&M\|\Pi_hg-g_h\|^2+\frac{1}{2}\frac{d}{dt}\|r_h(\varsigma)-r_h\|^2\\&\le M\|q_h(\varsigma)-q_h\|\|r_h(\varsigma)-r_h\|+M\|\varrho_h(\varsigma)-\varrho_h\|\|\Pi_hg-g_h\|\\&+M\|g_h(\varsigma)-g\|\|\Pi_hg-g_h\|+M\|g-\Pi_hg\|\|\Pi_hg-g_h\|\\&+Mh\|g\|_1\|\Pi_hg-g_h\|+M\|div(g_h(\varsigma)-g)\|\|r_h(\varsigma)-r_h\|,\end{aligned} \tag{3.24}$$

where we also used the inequality $\sigma(B^{-1}\Pi_hv,\pi)\le Mh\|v\|_1\|\pi\|$. Next,we use the ε-Cauchy inequality, (3.3) and Lemma 3.1 to estimate (3.24),

$$\begin{aligned}&M\|\Pi_hg-g_h\|^2+\frac{1}{2}\frac{d}{dt}\|r_h(\varsigma)-r_h\|^2\\&\le M\|q_h(\varsigma)-q_h\|^2+M\|\varrho_h(\varsigma)-\varrho_h\|^2\\&+\varepsilon\|\Pi_hg-g_h\|^2+\varepsilon\|r_h(\varsigma)-r_h\|^2+Mh^2.\end{aligned} \tag{3.25}$$

We integrate (3.25) over time as well as making use of the Gronwall's lemma, where we will use $(r_h(\varsigma)-r_h)|_{t=T}=0$. So we can obtain

$$\begin{aligned}&\|\Pi_hg-g_h\|_{L^2(J;L^2(\Omega))}+\|r_h(\varsigma)-r_h\|_{L^\infty(J;L^2(\Omega))}\\&\le M\|q_h(\varsigma)-q_h\|_{L^2(J;L^2(\Omega))}+M\|\varrho_h(\varsigma)-\varrho_h\|_{L^2(J;L^2(\Omega))}+Mh.\end{aligned} \tag{3.26}$$

Combining (3.26), (3.3), (3.7) and (3.8), we will derive

$$\begin{aligned}
&\|g_h(\varsigma) - g_h\|_{L^2(J;L^2(\Omega))} + \|r_h(\varsigma) - r_h\|_{L^\infty(J;L^2(\Omega))} \\
&= \|g_h(\varsigma) - g\|_{L^2(J;L^2(\Omega))} + \|g - \Pi_h g\|_{L^2(J;L^2(\Omega))} + \|\Pi_h g - g_h\|_{L^2(J;L^2(\Omega))} + \|r_h(\varsigma) - r_h\|_{L^\infty(J;L^2(\Omega))} \\
&\leq Mh + Mh + M\|q_h(\varsigma) - q_h\|_{L^2(J;L^2(\Omega))} + M\|\varrho_h(\varsigma) - \varrho_h\|_{L^2(J;L^2(\Omega))} + Mh \\
&\leq M\|\varsigma - \varsigma_h\|_{L^2(J;L^2(\Omega))} + Mh.
\end{aligned} \tag{3.27}$$

Thus, the theorem has been proofed. □

From reference [16], we get the following result.

Lemma 3.2. *For any $\tilde{\varsigma}_h \in K_h$, there exists a function $\tilde{\varsigma}^* \in K$, such that*

$$\tilde{\varsigma}^*_T = \tilde{\varsigma}_h|_T, \forall T \in \Gamma^h. \tag{3.28}$$

And some error estimates as follows:

$$\|\tilde{\varsigma}^*\|_{1,T} \leq (\|\alpha\|^2_{1,T} + \|\beta\|^2_{1,T})^{\frac{1}{2}}, \tag{3.29}$$

$$\|\tilde{\varsigma}_h - \tilde{\varsigma}^*\|_{-1} \leq Mh^2(\|\alpha\|^2_1 + \|\beta\|^2_1)^{\frac{1}{2}}. \tag{3.30}$$

Theorem 3.2. *Let ς and ς_h be the optimal controls of (2.4)–(2.8) and (2.15)–(2.19) respectively; for $M > 0$, which is a constant and independent of h; such that*

$$\|\varsigma - \varsigma_h\|_{L^2(J;L^2(\Omega))} \leq Mh, \tag{3.31}$$

$$\|\varrho - \varrho_h\|_{L^2(J;L^2(\Omega))} + \|q - q_h\|_{L^\infty(J;L^2(\Omega))} \leq Mh, \tag{3.32}$$

$$\|g - g_h\|_{L^2(J;L^2(\Omega))} + \|r - r_h\|_{L^\infty(J;L^2(\Omega))} \leq Mh. \tag{3.33}$$

Proof. First of all, let's do a little bit of preparation for estimating $\|\varsigma - \varsigma_h\|$. For $\tilde{\varsigma} \in K$, we use the multipoint method to discrete (2.20)–(2.23), then we have

$$(B^{-1}\varrho_h(\tilde{\varsigma}), \mathbf{v}_h)_Q - (q_h(\tilde{\varsigma}), div\mathbf{v}_h) = 0, \forall \mathbf{v}_h \in V_h, \tag{3.34}$$

$$(q_{ht}(\tilde{\varsigma}), \pi_h) + (div\varrho_h(\tilde{\varsigma}), \pi_h) = (\tilde{\varsigma}, \pi_h), \forall \pi_h \in W_h, \tag{3.35}$$

$$(B^{-1}g_h(\tilde{\varsigma}), \mathbf{v}_h)_Q - (r_h(\tilde{\varsigma}), div\mathbf{v}_h) = -(\varrho_h(\tilde{\varsigma}) - \varrho_d, \mathbf{v}_h), \forall \mathbf{v}_h \in V_h, \tag{3.36}$$

$$-(r_{ht}(\tilde{\varsigma}), \pi_h) + (divg_h(\tilde{\varsigma}), \pi_h) = (q_h(\tilde{\varsigma}) - q_d, \pi_h), \forall \pi_h \in W_h. \tag{3.37}$$

For any $\tilde{\varsigma} \in K$ and $\tilde{\varsigma}_h \in K_h$, from (2.16) and (3.35) with $\tilde{\varsigma} = \tilde{\varsigma}_h$ so that

$$(div(\varrho_h - \varrho_h(\tilde{\varsigma}_h), \pi_h) = (\varsigma_h - \tilde{\varsigma}_h, \pi_h) - (q_{ht} - q_{ht}(\tilde{\varsigma}_h), \pi_h), \tag{3.38}$$

making a difference between (2.5) and (3.35), we have

$$(div(\varrho - \varrho_h(\tilde{\varsigma}), \pi_h) = (\varsigma - \tilde{\varsigma}, \pi_h) - (q_t - q_{ht}(\tilde{\varsigma}), \pi_h). \tag{3.39}$$

It follows from $(r_h + \varsigma_h, \tilde{\varsigma}_h - \varsigma_h) \geq 0, \forall \tilde{\varsigma}_h \in K_h$ that

$$(\varsigma_h, \varsigma_h - \tilde{\varsigma}_h) + (r_h, \varsigma_h - \tilde{\varsigma}_h) \leq 0. \tag{3.40}$$

From (2.17) and (3.38), we have that

$$\begin{aligned}(\varsigma_h - \tilde{\varsigma}_h, r_h) &= (q_{ht} - q_{ht}(\tilde{\varsigma}_h), r_h) + (div(\varrho_h - \varrho_h(\tilde{\varsigma}_h)), r_h)\\ &= (q_{ht} - q_{ht}(\tilde{\varsigma}_h), r_h) + (B^{-1}g_h, \varrho_h - \varrho_h(\tilde{\varsigma}_h))_Q + (\varrho_h - \varrho_d, \varrho_h - \varrho_h(\tilde{\varsigma}_h)),\end{aligned} \tag{3.41}$$

from (2.15), (3.34), (2.18), we can obtain

$$\begin{aligned}&(B^{-1}g_h, \varrho_h - \varrho_h(\tilde{\varsigma}_h))_Q\\ &= (B^{-1}(\varrho_h - \varrho_h(\tilde{\varsigma}_h)), g_h)_Q\\ &= (q_h - q_h(\tilde{\varsigma}_h), divg_h)\\ &= (q_h - q_d, q_h - q_h(\tilde{\varsigma}_h)) + (r_{ht}, q_h - q_h(\tilde{\varsigma}_h)).\end{aligned} \tag{3.42}$$

Considering (3.40)–(3.42), we can derive

$$\begin{aligned}&(\varsigma_h, \varsigma_h - \tilde{\varsigma}_h) + (q_{ht} - q_{ht}(\tilde{\varsigma}_h), r_h) + (q_h - q_d, q_h - q_h(\tilde{\varsigma}_h)) + (r_{ht}, q_h - q_h(\tilde{\varsigma}_h))\\ &+(\varrho_h - \varrho_d, \varrho_h - \varrho_h(\tilde{\varsigma}_h)) \leq 0.\end{aligned} \tag{3.43}$$

Next, relations (2.17), (2.18), (3.34) and (3.35) imply that any $\tilde{\varsigma}_h \in K_h$,

$$\begin{aligned}&(\varrho_h - \varrho_d, \varrho_h(\tilde{\varsigma}_h)) + (q_h - q_d, q_h(\tilde{\varsigma}_h))\\ &= -(B^{-1}g_h, \varrho_h(\tilde{\varsigma}_h))_Q + (r_h, div\varrho_h(\tilde{\varsigma}_h)) - (r_{ht}, q_h(\tilde{\varsigma}_h)) + (divg_h, q_h(\tilde{\varsigma}_h))\\ &= (\tilde{\varsigma}_h, r_h) - (q_{ht}(\tilde{\varsigma}_h), r_h) - (r_{ht}, q_h(\tilde{\varsigma}_h)).\end{aligned} \tag{3.44}$$

Let $\tilde{\varsigma} = \varsigma$ in (2.17) and (2.18), (3.34) and (3.35), we can find that

$$(\varrho_h - \varrho_d, \varrho_h(\varsigma)) + (q_h - q_d, q_h(\varsigma)) = (\varsigma, r_h) - (q_{ht}(\varsigma), r_h) - (r_{ht}, q_h(\varsigma)). \tag{3.45}$$

From (3.44) and (3.45), we can get that

$$\begin{aligned}&(\varrho_h - \varrho_d, \varrho_h(\tilde{\varsigma}_h) - \varrho_h(\varsigma)) + (q_h - q_d, q_h(\tilde{\varsigma}_h) - q_h(\varsigma))\\ &= (\tilde{\varsigma}_h - \varsigma, r_h) - (q_{ht}(\tilde{\varsigma}_h) - q_{ht}(\varsigma), r_h) - (r_{ht}, q_h(\tilde{\varsigma}_h) - q_h(\varsigma)).\end{aligned} \tag{3.46}$$

Similarly, there is

$$\begin{aligned}&(\varrho_h(\varsigma) - \varrho_d, \varrho_h - \varrho_h(\varsigma)) + (q_h(\varsigma) - q_d, q_h - q_h(\varsigma))\\ &= (\varsigma_h - \varsigma, r_h(\varsigma)) - (q_{ht} - q_{ht}(\varsigma), r_h(\varsigma)) - (r_{ht}(\varsigma), q_h - q_h(\varsigma)).\end{aligned} \tag{3.47}$$

At present, we will use the above results to estimate $\|\varsigma - \varsigma_h\|$. We study that

$$\begin{aligned}\|\varsigma - \varsigma_h\|^2 &\leq (\varsigma_h, \varsigma_h - \varsigma) + (\varrho_h - \varrho_d, \varrho_h - \varrho_h(\varsigma)) + (q_h - q_d, q_h - q_h(\varsigma))\\ &-(\varsigma, \varsigma_h - \varsigma) - (\varrho_h(\varsigma) - \varrho_d, \varrho_h - \varrho_h(\varsigma)) - (q_h(\varsigma) - q_d, q_h - q_h(\varsigma))\\ &= (\varsigma_h, \varsigma_h - \tilde{\varsigma}_h) + (\varrho_h - \varrho_d, \varrho_h - \varrho_h(\tilde{\varsigma}_h)) + (q_h - q_d, q_h - q_h(\tilde{\varsigma}_h))\\ &+(\varrho_h - \varrho_d, \varrho_h(\tilde{\varsigma}_h) - \varrho_h(\varsigma)) + (q_h - q_d, q_h(\tilde{\varsigma}_h) - q_h(\varsigma)) - (\varrho_h(\varsigma) - \varrho_d, \varrho_h - \varrho_h(\varsigma))\\ &-(q_h(\varsigma) - q_d, q_h - q_h(\varsigma)) + (\varsigma_h, \tilde{\varsigma}_h - \varsigma) - (\varsigma, \varsigma_h - \varsigma).\end{aligned} \tag{3.48}$$

Applying (3.43), (3.46) and (3.47), there is

$$\begin{aligned}
&\|\varsigma-\varsigma_h\|^2 \le -(q_{ht}-q_{ht}(\tilde{\varsigma}_h),r_h)-(r_{ht},q_h-q_h(\tilde{\varsigma}_h))+(\tilde{\varsigma}_h-\varsigma,r_h)-(q_{ht}(\tilde{\varsigma}_h)-q_{ht}(\varsigma),r_h)\\
&-(r_{ht},q_h(\tilde{\varsigma}_h)-q_h(\varsigma))-(\varsigma_h-\varsigma,r_h(\varsigma))+(q_{ht}-q_{ht}(\varsigma),r_h(\varsigma))+(r_{ht}(\varsigma),q_h-q_h(\varsigma))\\
&+(\varsigma_h,\tilde{\varsigma}_h-\varsigma)-(\varsigma,\varsigma_h-\varsigma)\\
&=-(q_{ht}-q_{ht}(\varsigma),r_h)-(r_{ht},q_h-q_h(\varsigma))+(\tilde{\varsigma}_h-\varsigma,\varsigma_h+r_h)-(\varsigma_h-\varsigma,\varsigma+r_h(\varsigma))\\
&+(q_{ht}-q_{ht}(\varsigma),r_h(\varsigma))+(r_{ht}(\varsigma),q_h-q_h(\varsigma))\\
&=-(q_{ht}-q_{ht}(\varsigma),r_h-r_h(\varsigma))-(r_{ht}-r_{ht}(\varsigma),q_h-q_h(\varsigma))+(\tilde{\varsigma}_h-\varsigma,\varsigma_h+r_h)\\
&-(\varsigma_h-\varsigma,\varsigma+r_h(\varsigma)).
\end{aligned} \tag{3.49}$$

We know that

$$(\varsigma+r,\tilde{\varsigma}-\varsigma)\ge 0. \tag{3.50}$$

Hence, from (3.50), we will get

$$(\varsigma+r,\tilde{\varsigma}-\varsigma_h)+(\varsigma+r_h(\varsigma),\varsigma_h-\varsigma)+(r-r_h(\varsigma),\varsigma_h-\varsigma)\ge 0. \tag{3.51}$$

Thus, we can get $\|\varsigma-\varsigma_h\|^2$ with (3.49)–(3.51) as follows

$$\begin{aligned}
&\|\varsigma-\varsigma_h\|^2 \le -(q_{ht}-q_{ht}(\varsigma),r_h-r_h(\varsigma))-(r_{ht}-r_{ht}(\varsigma),q_h-q_h(\varsigma))\\
&+(\tilde{\varsigma}_h-\varsigma,\varsigma_h+r_h)+(\varsigma+r,\tilde{\varsigma}-\varsigma_h)+(r-r_h(\varsigma),\varsigma_h-\varsigma)\\
&\le -(q_{ht}-q_{ht}(\varsigma),r_h-r_h(\varsigma))-(r_{ht}-r_{ht}(\varsigma),q_h-q_h(\varsigma))\\
&+(\varsigma+r,\tilde{\varsigma}_h-\varsigma)+(\varsigma_h-\varsigma,\tilde{\varsigma}_h-\varsigma)-(r-r_h(\varsigma),\tilde{\varsigma}_h-\varsigma)\\
&+(r_h-r_h(\varsigma),\tilde{\varsigma}_h-\varsigma)+(\varsigma+r,\tilde{\varsigma}-\varsigma_h)+(r-r_h(\varsigma),\varsigma_h-\varsigma)\\
&\le -\frac{d}{dt}(q_h-q_h(\varsigma),r_h-r_h(\varsigma))+\|\varsigma+r\|_1(\|\tilde{\varsigma}-\varsigma_h\|_{-1}+\|\tilde{\varsigma}_h-\varsigma\|_{-1})\\
&+(\|\varsigma-\varsigma_h\|+\|r_h-r_h(\varsigma)\|)\|\tilde{\varsigma}_h-\varsigma\|+\|r-r_h(\varsigma)\|(\|\tilde{\varsigma}_h-\varsigma\|+\|\varsigma-\varsigma_h\|).
\end{aligned} \tag{3.52}$$

According to the Lemma 3.3, there exists $\varsigma^*\in K$ and for all $T\in\Gamma_h$, there are

$$(\varsigma^*)_T=\varsigma_h|_T,$$

$$\|\varsigma_h-\varsigma^*\|_{-1}\le Mh^2(\|\alpha\|_1^2+\|\beta\|_1^2)^{\frac{1}{2}}. \tag{3.53}$$

We choose $\tilde{\varsigma}=\varsigma^*$ and $\tilde{\varsigma}_h=P_h\varsigma$ in (3.52), combining (3.52) and (3.53), using (3.5). We also apply the ε-Cauchy inequality. Then we can obtained that

$$\begin{aligned}
&\|\varsigma-\varsigma_h\|^2 \le -\frac{d}{dt}(q_h-q_h(\varsigma),r_h-r_h(\varsigma))+Mh^2\|\varsigma+r\|_1\\
&+Mh(\|\varsigma-\varsigma_h\|+\|r_h-r_h(\varsigma)\|+\|r-r_h(\varsigma)\|)+\|r-r_h(\varsigma)\|\;\|\varsigma-\varsigma_h\|\\
&\le -\frac{d}{dt}(q_h-q_h(\varsigma),r_h-r_h(\varsigma))+Mh^2+\xi_1\|r_h-r_h(\varsigma)\|^2+M\|r-r_h(\varsigma)\|^2+\xi_2\|\varsigma-\varsigma_h\|^2,
\end{aligned} \tag{3.54}$$

for any small $\xi_1,\xi_2>0$. We know that,

$$(q_h-q_h(\varsigma))|_{t=0}=0,$$

$$(r_h-r_h(\varsigma))|_{t=T}=0.$$

Thus, we can easily get that

$$\int_0^T -\frac{d}{dt}(q_h - q_h(\varsigma), r_h - r_h(\varsigma)) = 0.$$

Then, integrating (3.54) in time, using Lemma 3.1, (3.8) and (3.9), we have

$$\|\varsigma - \varsigma_h\|_{L^2(J;L^2(\Omega))} \le Mh. \tag{3.55}$$

Then, we can get our final results:

$$\|\varrho - \varrho_h\|_{L^2(J;L^2(\Omega))} + \|q - q_h\|_{L^\infty(J;L^2(\Omega))} \le Mh, \tag{3.56}$$

$$\|g - g_h\|_{L^2(J;L^2(\Omega))} + \|r - r_h\|_{L^\infty(J;L^2(\Omega))} \le Mh. \tag{3.57}$$

So far, we have completed the proof of the theorem. □

4. Numerical experiments

In this section, we will verify the error estimates of the state (ϱ,q), co-state (g,r) and the control variable ς of theoretical analysis by a numerical example.

First, we consider the following OCP for parabolic equations:

$$\min_{\varsigma\in K}\left\{\frac{1}{2}\int_0^T (\|\varrho - \varrho_d\|^2 + \|q - q_d\|^2 + \|\varsigma\|^2)dt\right\}, \tag{4.1}$$

subject to the state equation and boundary conditions

$$\begin{cases} q_t + div\varrho = \varsigma + f, \varrho = -\nabla q, x \in \Omega, \\ q(x,t) = 0, x \in \partial\Omega, t \in J, \\ q(x,0) = 0, x \in \Omega, \\ -r_t + divg = q - q_d, g = -\nabla r - \varrho + \varrho_d, x \in \Omega, \\ r(x,t) = 0, x \in \partial\Omega, t \in J, \\ r(x,T) = 0, x \in \Omega. \end{cases} \tag{4.2}$$

We choose the domain $\Omega = [0,1]\times[0,1]$, $T = 1$. We adopt the same mesh partition for the state and the control. The convergence order is computed by the following formula: $order \simeq \frac{log(E_i/E_i+1)}{log(h_i/h_{i+1})}$, where i responds to the spatial partition, and E_i denote the L^2 norm for the state and co-state and the control approximation.

Example. The data for the numerical example is as follows:

$$\begin{aligned} &\varsigma = max(0.4 - r, 0), \\ &q = sin(2\pi x_1)sin(2\pi x_2)sin(\pi t), \\ &r = sin(2\pi x_1)sin(2\pi x_2)sin(\pi t), \\ &f = q_t + div\varrho - \varsigma, \\ &q_d = q + r_t, \\ &\varrho = (2\pi cos(2\pi x_1)sin(2\pi x_2)sin(\pi t), 2\pi sin(2\pi x_1)cos(2\pi x_2)sin(\pi t)), \\ &\varrho_d = (0,0), \\ &g = (0,0). \end{aligned} \tag{4.3}$$

Table 1. Numerical error for the state, co-state, and control variables.

subdivision	8×8	16×16	32×32	64×64
$\|\varrho-\varrho_h\|_{L^2(0,T;L^2(\Omega))}$	1.0539E-01	5.2170E-02	2.5770E-02	1.7195E-03
order	-	1.0144	1.0175	1.0123
$\|q-q_h\|_{L^2(0,T;L^2(\Omega))}$	1.4965E-02	7.5657E-03	3.7883E-03	1.8935E-03
order		0.9840	0.9979	1.0005
$\|g-g_h\|_{L^2(0,T;L^2(\Omega))}$	1.1926E-01	6.7083E-02	3.4607E-02	1.7405E-02
order	-	0.8301	0.9549	0.9916
$\|r-r_h\|_{L^2(0,T;L^2(\Omega))}$	1.3311E-02	6.7998E-03	3.4283E-03	1.7195E-03
order	-	0.9691	0.9880	0.9955
$\|\varsigma-\varsigma_h\|_{L^2(0,T;L^2(\Omega))}$	1.1583E-02	6.3890E-03	3.2607E-02	1.7138E-03
order	-	0.8580	0.9589	0.9808

In this example, the L^2 errors of $\|\varrho-\varrho_h\|_{L^2(0,T;L^2(\Omega))}$, $\|q-q_h\|_{L^2(0,T;L^2(\Omega))}$, $\|g-g_h\|_{L^2(0,T;L^2(\Omega))}$, $\|r-r_h\|_{L^2(0,T;L^2(\Omega))}$, $\|\varsigma-\varsigma_h\|_{L^2(0,T;L^2(\Omega))}$ on the MFMFE approximation for state functions and piecewise constant approximation for control function are presented in Table 1. Numerical results show that the grid subdivision is smaller, the smaller the error of state, co-state and control variables. And we can see that the order of state, co-state and control variables basically reaches first order, which is consistent with our theoretical analysis. So we can conclude from the table that the theoretical analysis is correct.

5. Conclusions

We innovatively use the semi-discrete MFMFE method to study the error for the state, co-state and the control varibles of Parabolic OCP. Our method is a decoupled method, which avoids the problem of solving saddle-point algebraic systems required by MFE. It is very novel in dealing with parabolic optimal control problems, and our method also gets the same results as others after dealing with the problems, which further proves the correctness of our method and theory. Finally, we also verify the theoretical analysis through a numerical example.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (11801293), the Natural Science Foundation of Shandong Province (ZR2020MA049) and the Education and Industry Integration Pilot Project Basic Research Project of Qilu University of Technology (Shandong Academy of Sciences) (2022PY058).

Conflict of interest

This work does not have any conflict of interest.

References

1. W. Alt, On the approximation of infinite optimization problems with an application to optimal control problems, *Appl. Math. Optim.*, **12** (1984), 15–27. https://doi.org/10.1007/BF01449031

2. F. Falk, Approximation of a class of optimal control problems with order of convergence estimates, *J. Math. Anal. Appl.*, **44** (1973), 28–47. https://doi.org/10.1016/0022-247X(73)90022-X

3. K. Malanowski, Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal-control problems, *Appl. Math. Optim.*, **8** (1982), 69–95. https://doi.org/10.1007/BF01447752

4. M. Yan, W. Gong, N. Yan, Finite element methods for elliptic optimal control problems with boundary observations, *Appl. Numer. Math.*, **90** (2015), 190–207. http://dx.doi.org/10.1016/j.apnum.2014.11.011

5. T. Hou, C. Liu, Y. Yang, Error estimates and superconvergence of a mixed finite element method for elliptic optimal control problems, *Comput. Math. Appl.*, **74** (2017), 714–726. http://dx.doi.org/10.1016/j.camwa.2017.05.021

6. H. Choi, W. Choi, Y. Koh, A finite element method for elliptic optimal control problem on a non-convex polygon with corner singularities, *Comput. Math. Appl.*, **75** (2018), 45–58. http://dx.doi.org/10.1016/j.camwa.2017.08.029

7. Z. Zhang, D. Liang, Q. Wang, Immersed finite element method and its analysis for parabolic optimal control problems with interfaces, *Appl. Numer. Math.*, **147** (2020), 174–195. https://doi.org/10.1016/j.apnum.2019.08.024

8. S. Brenner, M. Oh, L. Sung, P_1 finite element methods for an elliptic state-constrained distributed optimal control problem with Neumann boundary conditions, *Results in Applied Mathematics*, **8** (2020), 100090. https://doi.org/10.1016/j.rinam.2019.100090

9. K. Porwal, P. Shakya, A finite element method for an elliptic optimal control problem with integral state constraints, *Appl. Numer. Math.*, **169** (2021), 273–288. https://doi.org/10.1016/j.apnum.2021.07.002

10. P. Neittaanmaki, D. Tiba, *Optimal control of nonlinear parabolic systems: theory, algorithms and applications*, New York: Marcel Dekker Press, 1994.

11. W. Liu, N. Yan, *Adaptive finite element methods for optimal control governed by PDEs*, Beijing: Science Press, 2008.

12. J. Lions, *Optimal control of systems governed by partial differential equations*, Berlin: Springer-Verlag Press, 1971.

13. H. Guo, H. Fu, J. Zhang, A splitting positive definite mixed finite element method for elliptic optimal control problem, *Appl. Math. Comput.*, **219** (2013), 11178–11190. https://doi.org/10.1016/j.amc.2013.05.020

14. W. Liu, N. Yan, A posteriori error estimates for distributed convex optimal control problems, *Adv. Comput. Math.*, **15** (2001), 285–309. https://doi.org/10.1023/A:1014239012739

15. P. Shakya, R. Sinha, Finite element method for parabolic optimal control problems with a bilinear state equation, *J. Comput. Appl. Math.*, **367** (2020), 112431. https://doi.org/10.1016/j.cam.2019.112431

16. X. Xing, Y. Chen, Error estimates of mixed methods for optimal control problem by parabolic equations, *Int. J. Numer. Meth. Eng.*, **75** (2008), 735–754. https://doi.org/10.1002/nme.2289

17. Y. Chen, Z. Lu, Error estimates for parabolic optimal control problem by fully discrete mixed finite element methods, *Finite Elem. Anal. Des.*, **46** (2010), 957–965. https://doi.org/10.1016/j.finel.2010.06.011

18. I. Aavatsmark, An introduction to multipoint flux approximations for quadrilateral grids, *Computat. Geosci.*, **6** (2002), 405–432. https://doi.org/10.1023/A:1021291114475

19. T. Arbogast, C. Dawson, P. Keenan, M. Wheeler, I. Yotov, Enhanced cell-centered finite differences for elliptic equations on general geometry, *SIAM J. Sci. Comput.*, **19** (1998), 404–425. https://doi.org/10.1137/S1064827594264545

20. M. Wheeler, I. Yotov, A multipoint flux mixed finite element method, *SIAM J. Numer. Anal.*, **44** (2006), 2082–2106. https://doi.org/10.1137/050638473

21. W. Xu, Cell-centered finite difference method for parabolic equation, *Appl. Math. Comput.*, **235** (2014), 66–79. http://dx.doi.org/10.1016/j.amc.2014.02.066